Journal of Forestry Research (2008) 19(4):288–292
DOI 10.1007/s11676-008-0051-2

RESEARCH PAPER

Analysis and identification of SCAR molecular markers associated with birch fiber length trait

WANG Dan, WEI Zhi-gang*, YANG Chuan-ping, LIU Guan-jun

Key Laboratory of Forest Tree Genetic and Biotechnology, Ministry of Education, Northeast Forestry University, Heilongjiang, Haerbin, 150040, China

Abstract: The fiber length trait (FLT) of 538 individuals from nature birch population in Maorshan region, Heilongjang, China were measured, of which 100 individuals were selected as representative variety of correlated fragments screening with random amplified polymorphism DNA (RAPD) technique. In total of 20 RAPD primers were tested through multiple regression analysis between amplified strip and the character behaviors, and a correlative segment BFLR-16 was obtained. The correlation coefficient between BFLI-16 and FLT was 0.6144, with the significant level of 1%. BFLI-16 was then cloned, sequenced and transformed into SCAR marker. The percentage of identifying long fiber birches by this SCAR was more than 92. The result indicates that the SCAR markers has high specificity for the long fiber individuals and is highly linked with the gene controlling the character of fiber length, and its existence is significantly correlative with the increase in the fiber length.

Keywords: *Betula platyphylla*; fiber length trait; random amplified polymorphism DNA; sequence characterized amplified region

Introduction

Betula platyphylla (silver birch) has the characteristics of well updated, rapid growth, strong adaptability, wide distributing, high content of cellulose with low content of lignin and white color etc. (Li et al. 1995). It is recognized as a kind of superordinary and short-period material for paper pulp. Studies showed that the fiber length was an important index for the mechanical strength of paper, which affects paper's qualities. Therefore, one of important aims for researching the fiber length trait (FTL) of birch is to select new species of birch. Molecular labeling technique has an incomparable advantages compared with traditional breeding methods, and it has been widely used to establish genetic linkage maps of trees and quantitative trait locus (QTLs) (Cervera et al. 2001; Wu et al. 2000). However, the QTLs of a special trait selected from a single-population is greatly limited due to the restriction of labeling theories and technical methods. And the comparability between the results of the QTLs of different labs selected from different populations is low (Zhang et al. 2003; Grattapaglia et al. 1994; Jermstad et al. 2003). The researched populations and analysis methods of QTLs need improvement. The main random molecular labeling techniques used in the research of QTLs such as RAPD (Random amplification polymorphism DNA), ISSR (Inter-simple sequence repeat) and AFLP (Amplified fragment length polymorphism) have the disadvantages of low specificity and stability, and the domain amplified is unknown. Therefore, the gained QTLs can hardly be used in trees' labeling aided to selecting breeding. In the SCAR labeling technique, a pair of specific primers is designed according to a known sequence of the genome. Then a restriction region of the genome is amplified with high specificity, which is used as a specific labeling to the genome. This method overcomes the limitation of the RAPD markers technique and is widely used for locating relative genes of a correlative traits and establishing genetic maps (Meletto 1996; Garcia 1996; Ohmori 1996).

In this study, a fragment closely correlated with fiber length trail of birch was selected using the random amplified polymorphism DNA (RAPD) technique, then sequenced and cloned, finally transformed into a stable SCAR marker. The aim of this study is 1) to provide a tool for high accuracy marker assistant selection in longer fiber birch, and 2) to set a good background and a new method for marker assistant breeding on other species.

Materials and methods

Plant materials

A total of 583 individuals of silver birch were randomly selected from natural birch forest in Maorshan Forest Experiment Farm of Northeast Forestry University, Heilongjiang, China. The core

Foundation project: This study was supported by the National 863 Program (2002BA515B0401), National Natural Science Foundation of China (30571513) and Foundation of Heilongjiang Province (GA06B301)

Received: 2008-05-29; Accepted: 2008-07-03

The online version is available at http://www.springerlink.com

Biography: WANG Dan (1982-), female, postgraduate student studies at Northeast Forestry University, interested in forest tree genetics and biotechnology. Email: wangdanwana@163.com

*Corresponding author: WEI Zhigang (Email: zhigangwe@163.com)

Responsible editor: Hu Yanbo

was collected with an awl at diameter at breast height (DBH) from south to north. Tender leaves were collected with scissors at the same time. The leaves and the core were ensured to be from the same tree in order to make markers. The collected leaves were kept immediately in the refrigerator at -20°C. To identify the validity of the segments, the verified materials of 200 birch individuals selected randomly were collected at the Harbin Experimental Forestry Farm of Northeast Forestry University with the same methods. These 200 individuals did not have any genetic relationship with the natural birch forest in Maorshan.

Methods

The determination of fiber length

The pith was cut into small pieces with the same length and thickness as matchsticks. Then it was kept in 30%–50% nitric acid with slight potassium chloride solution. This can decompose the pith into cellulose. The fiber length was measured under optical microscope with an enlarged multiple of 40 to 70. In each sample, 10 pieces of fiber selected randomly were measured and the average value standed for the firer length of this sample.

Genomic DNA extracting

The improved CTAB methods were used to extract the whole DNA of the silver birch (Zhan et al. 2005).

Analysis of RAPD

The RAPD reaction system was determined through orthogonal design methods, on which the best reaction system of silver birch was gained. The 20-μL reaction system includes 2.4-μL Mg^{2+} (25 mmol/L), 0.8-μL dNTPs (2.5 mmol/L), 1.5-μL primer (50 pmol/μL), 0.1-μL Taq DNA polymerase (5 U/μL), and 2.5-μL template DNA (15 ng/μL). Amplification reaction was performed on a thermal cycler (MJ-9600PCR), programmed as follows: 3 min at 94°C for initial denaturizing, then 38 cycles of 94°C for 30 s, 38°C for 30 s, 72°C for 90 s, and extended at 72°C for 7 min, finally keeping it at 4°C.

The individuals with the longest and the shortest fiber length were taken as samples and 20 RAPD primers were screened for good polymorphism. 100 genomic DNA of the silver birch was used as template for the amplification reaction, and PCR products were separated by electrophoresis at 150 V for 2 h, through the agar gel of 1.5%. After taking pictures in the UVP gel imaging system, the diverse strip was noted, 1 stands for having strip and 0 for none.

The calculation of correlation coefficient

Taking the genotypic value of signed segments as the independent variable and the fiber length of 100 selected birch individuals as variable, we establish the multiple regression equation as follows:

$$y_n = \mu + b_{11}x_{11} + b_{12}x_{12} + k + \ldots b_{ni}x_{ni}$$

Where, y_n is the fiber length of birch individual (n = 1,2……100), μ the regression constant, b_{ni} the correlation coefficient of signed segment with fiber length (i is the sequence number of the signed segments, i = 1,2,3,4,5,6,7……), and X_{ni} represents the genotypic value of a signed segment of the silver birch having the sequence number of n (X, 0/1).

The relativity of fiber length with molecular markers is a respective multiple regression coefficient. When the regression coefficient is positive, it indicates the existence of signed segments is positively correlated with the increase in fiber length. When negative it indicates the existence of signed segments is negatively correlated with the decrease in fiber length. The significance level of the regression coefficient represents the degree of gene controlling the character of fiber length linked with the signed segments. If the significance level is high, it indicates the degree of gene linked is tight. Using the software package SPSS 12.0, we analyzed the regression coefficient of all signed segments, from which we selected the signed segments with high significance level. Those segments were reclaimed and transformed to SCAR markers.

Reclaiming and sequencing the segments with a high significance level

PCR products were separated by electrophoresis on the agar gel of 1.5%. The strip notably correlative with fiber length was cut from the gel. Reclaimed it with the reclaiming reagent boxes (purchased from bao biological company of dalian). It was cloned to the carrier of PMD18-T and transformed to the TOP10 strain. After checking with PCR, it was sent to be sequenced by bao biological company of dalian.

Transformation and verification of SCAR markers

Suitable primers were designed with software package primer 5.0 according to the results of sequencing. The primers were synthesized by shanghai biological company. The 20-μL amplified system includes 1.2-μL Mg^{2+} (25 mmol/L), 0.4-μL dNTPs (2.5 mmol/L), 0.4-μL cis-primer (50 pmol/μL), 0.4-μL trans-primer (50 pmol/μL), 1-μL Taq DNA polymerase (1 U/μL), 1-μL template DNA (40 ng/μL), and 2-μL 10 × PCR buffer. Amplification reaction was performed on a thermal cycler (MJ-9600PCR) programmed as follows: 4 min at 94°C for initial denaturizing, then 35 cycles of 94°C for 45 s, 56°C for 45 s,72°C for 80 s, and extended at 72°C for 7 min, finally keeping it at 4°C. PCR products were separated by electrophoresis at 120 V for about 2 h, through the agar gel of 2%. And the pictures were taken in the UVP gel imaging system.

Results and analysis

Selection of birch individuals

The fiber length of 583 individuals was analyzed with the software package of SPSS. Results showed that the average fiber length of the 583 individual was 4.73, which exhibited a typical normal distribution (Fig. 1). The results were in consistence with the feature of phenotype distribution of quantity genes control-

ling characters. If the fiber length is larger than the mean value (4.73), then a plus standard deviation (0.36) is given; oppositely, minus standard deviation is given. Base on the rule, we respectively chose 50 individuals with the longest and shortest fiber length to do the genome's discrepancy analysis. The fiber length of 200 individuals selected from the Harbin Experimental Forest Farm of Northeast Forestry University was analyzed with the software package of SPSS. Results showed that the average fiber length was 5.46 and the population showed a normal distribution (Fig. 2). The results were consistent with the feature of phenotype distribution of quantity gene controlling characters. Based on the above rule (the mean value (5.46), the standard deviation (0.60)), we selected 25 individuals with the longest fiber length and 25 individuals with the shortest as the verified population.

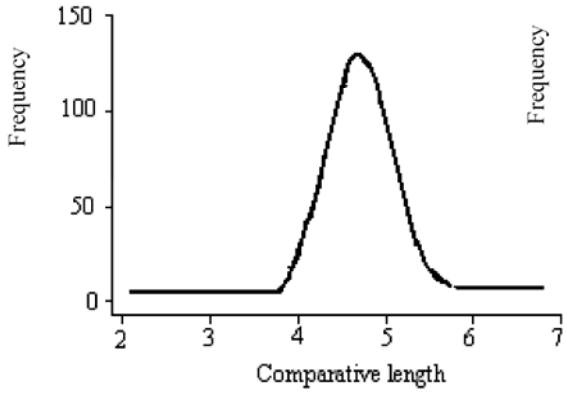

Fig. 1 Distribution of fiber length of 583 birches

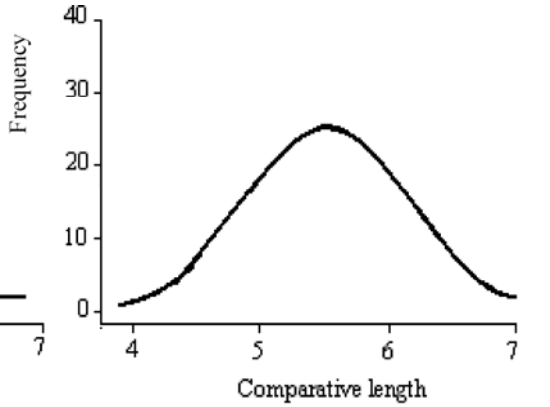

Fig. 2 Distribution of fiber length of 100 birches

Screening of relative RAPD signed segments

A total of 20 RAPD primers were used to analyze the genome of 100 birch individuals selected. Each primer had 0 to 1 diverse strip. The primer R-16 (TGCCGAGCTG) gave one different strip and the correlation coefficient was 0.6144. This segment was named "BFLR-16" with a molecular weight of 1200 bp or so (Fig. 3).

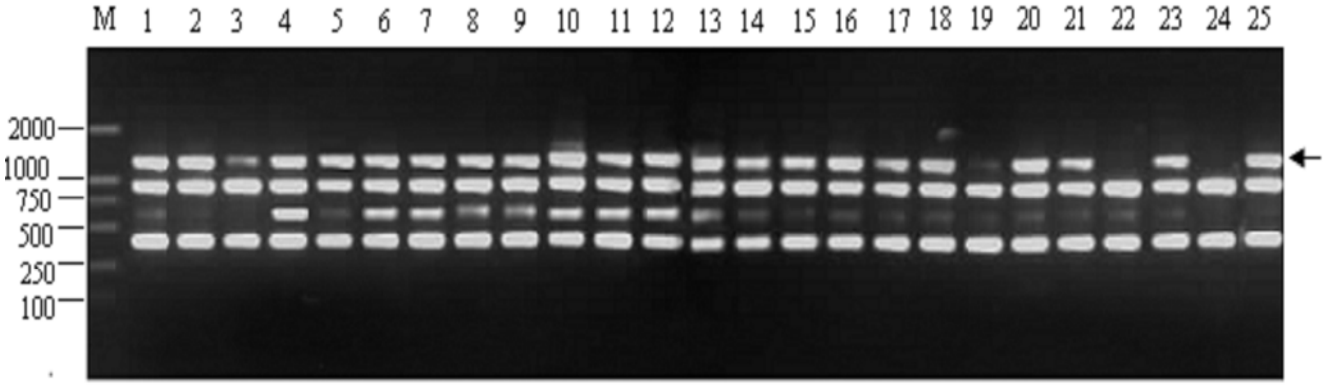

Fig. 3 Electrophoresis of amplified products with primer R-16 in some long fiber individuals M, DL2000 DNA molecular marker. 1-25: Long fiber individuals; the arrow indicates the correlated fragment

The reclamation and sequence of the correlation segments

The correlative segment BFLR-16 was reclaimed. The amplified reaction was conducted again with this segment as template and R-16 as primer. The molecular weight of the amplified segment was about 1200 bp (Fig. 4), matched with the objective segment, therefore, can be used to conjunction and clone directly. The reclaimed segments were connected to the carrier of PMD18-T and then transformed to the TOP10 strain. After cultured in the plate, 3 to 4 single white colonies were selected randomly to be cultured in the rocking bed. The plasmids were extracted. The amplified reaction was conducted using R-16 as primer to detect the correctness of the inserted segments (Fig. 5). Results indicated the recombinant plasmid had only one strip after electrophoresis and the molecular weight was consistent with correlative RAPD segments with long fiber trait. It demonstrates the RAPD specific segment has been correctively inserted into the plasmid. The bacterium including RAPD specific segment correlative with long fiber trait was sent to bao biological company to be sequenced. The sequence was as follow:

TGCCGAGCTGCTCAAATCCATTAATCCCAATCCATGATAATGGGTAAATTGGGTAACAATTGCATTGTTTAAAAAAACTCCC
AAAATTAACACGAAGTGTCGTCGTGGTGCAAACTCCCAAATGATAATTAA**GGTGAACTGTCATTTGTACC**AACAAATGTT
TCTAGGGAAAAAGAAAATAAACACATAAATTCACGACTTAATAAGTAATATCAGAATAAATACAACAGAATCAATCTAAGTT
AAGTAATATTTCTTTCTCAAAAATCAACTTCATTAGAAAAAAAACTATTTACCTTGATAACTTTACAAGTTAAAGCCTCCAG
AATGCATTAACATCTAAAATATACTATTCTAGCATAAATACAACAGTACCCAAAAAATTCATGTAAATCTGCTCTTCAGGACT
TTAGCCAGCAACTCAAAACCTCACATTAACTCTAATTTGCAATTTGCTAATATGTAAAACAATAAATTGCCAACGAAATTTA
GCCCATGCAGTGAGCTCGGTTCGCATAAATACAACAGTACCCATGTTGGAAACAAAATATCATTAAGGCTTAGTTTACAGTT
CATGTAAACTAATTTGCTAATATGTAATTCTTATTATGAGAAAAACTAGAGTGGGAATGAAATTTATCAGTTGTAAAGCTTATT
TCTTCACCACACCTAAATTTTATCCAGAATGCATTAACACCCAAAAAATATCTAAAATATACTATTCTAAAAACCCCGAGGCT
AAATTCATGTAAATCTGCTCTTCAGGACTTTAGCCAGCAACTCAAAACCTCACATTAACTCTAATTTGCAACCCCGAGGCTA
ATTTGCTAATATGTAAAACAATAATGTTTCTAGAAAATAAACACATAAATTCACGACTTAAAAGCCTCCAGAATGCATTAACA
TCTAAAATATACTATTCTAGCATAAATACAAACAATAATGTTTCTAGGGAAAAAGAAAATAAATCTGCTCTT**GACCACAATA**
GGCAGCTTCACCGAGGCTAGAGTGGGAATGAAATTAGTAACCTCACATTAACTCTAATATCAGTTGTAAAGCTTATTCACA
CCTAAATTACATTAACTCTATCTTCACCACACCTAAATTACTCTAATTACATTATTATGGTATTCAAGTAACCTCACATTAACTC
TAATTACATTAACTCTATGCAACCTTAT**CAGCTCGGCA**

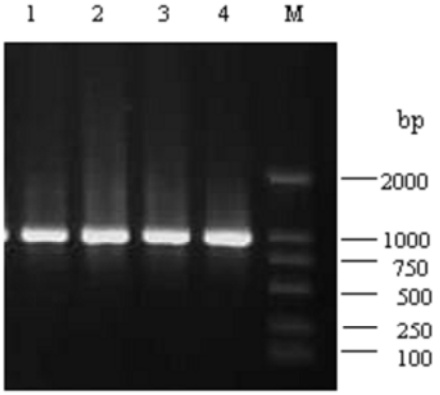

Fig. 4 Electrophoresis of amplification products with primer R-16 for callback M, DL2000 DNA molecular marker 1-4, Electrophoresis of amplification products for callback

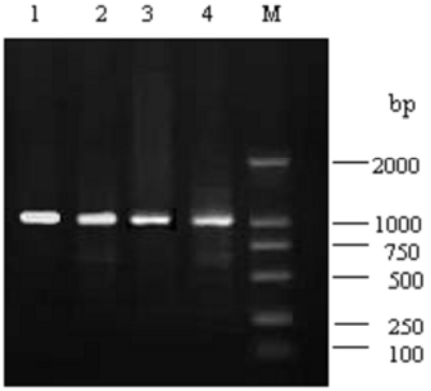

Fig. 5 Electrophoresis of amplification products with primer R-16 for recombines TOP10 M, DL2000 DNA molecular marker 1-4, Electrophoresis of amplification products for recombines TOP10

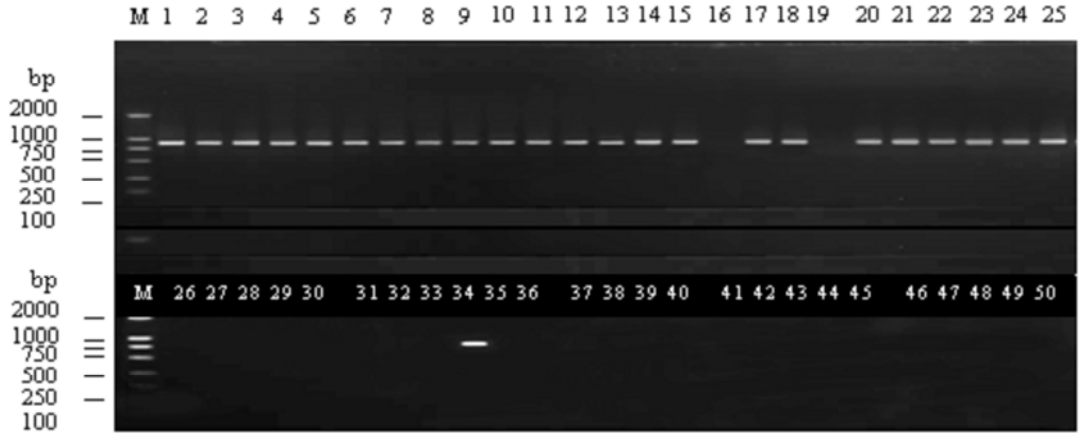

Fig. 6 Electrophoresis of the amplified products with primer BFL i n 50 individuals birch Note: M, DL2000 DNA molecular marker: 1-25, long fiber individuals, 26-50, short fiber individuals; The arrow indicates the correlated fragment

The verification of SCAR marker

According to the sequenced results, the specific primers with 20 bp were selected with the software of primer premier 5.0. This RAPD marker was transformed to SCAR marker. The sequence of BFLR-16 transformed to SCAR-R16 was as follow: BFLR-3F;

5′GGTGAACTGTCATTTGTACC3′; 5′TGAAGCTGCCTATTGTGGTC′. The length of the amplified segment was 440 bp. The genome of 50 birch individuals selected was used as template to be conducted the PCR with the designed primer. PCR products were separated by electrophoresis on the agar gel of 2%. There were 23 out of 25 individuals with longer fiber length having diverse strip (Fig. 6). The probability of long fiber individuals amplified was 92%. One out of 25 short fiber individuals was amplified. Therefore, it indicated the SCAR markers had high specificity for the long fiber individuals and may be highly linked with the gene controlling the character of fiber length.

Discussion

The forest has many important economic traits such as tree height, diameter at breast height, leaf area, and fiber length that are all quantitative characters (Zhang et al. 2002). These characters are controlled by many chromosomal loci (QTLs), usually distributed in different provenance sources or family constellations. The segregation population used to locate QTLs comes from the cross of two inbred lines, but this kind of segregation population is hardly to be obtained for perennial trees. Even if we can use the F1 progenies to substitute the segregation population, only part of the controlled characters can be identified (Zeng et al. 2006). This research gained 583 birch individuals with unknown genetic background. After doing multiple correlation analysis about the genotype and phenotypic characters of 100 typical individuals, we obtained the segments correlative with long fiber. Not only did we avoid the restriction of obtaining the QTLs hard for a single-population, but also there was hope to reveal all the QTLs systematically correlative with birch's long fiber trait as many character genes were gathered in the natural birch forest with complex genetic background.

The analytical method of correlation coefficient can be used to analyze and identify the sites of quantitative characters (Xu et al. 1994). The molecular marker correlative with the coefficient obtained by the method was consistent the results of variance analysis (Xu and Zou 2002). Compared with the mapping software MAPMAKER/QTL based on the maximum likelihood statistical single interval graphic method, the analytical method of correlation coefficient can be used to do correlation analysis about any solitary molecular marker. Using this method, there is no need to do several molecular markers to establish detailed and complete genetic maps and the material of birch can be both a specific segregation population and a combination of randomly selected individuals. This is helpful for analyzing systematically the expression state of the quantitative characters. Therefore, the method connecting the character of birch's fiber length with corresponding genomic marker is feasible. The markers correlated with birch's long fiber trait were obtained by multiple regression analysis. Also, it is possible to breakthrough the situation of retarded progression on marker aided selective breeding.

The RAPD method is easy to operate, but it is random marker and has the lowest rate of repeatability. Therefore, it's necessary to transform the obtained RAPD signed segments to the stable SCAR markers. In this research, we successfully transformed a RAPD segment "BFLR-16" to SCAR marker with an identification efficiency of 92% for long fiber birch, which can be used to aid selective breeding of long fiber birch. But the identification efficiency of SCAR marker is not absolute (Tang et al. 2005; Shi et al. 2001). Therefore, there is still need to make systematically research about every step of the process RAPD segments transforming to SCAR.

References

Cervera MT, Storme V, Ivens B, Gusmao J, Liu BH, Hostyn V, van Slycken J, van Montagu M, Boerjan W. 2001. Dense genetic linkage maps of three *populus* species (*Populus deltoides, P. nigra* and *P. trichocarpa*) based on AFLP and microsatellite markers. *Genetics*, **158**(2): 787–809.

Garcia GM. 1996. Identification of RAPD, SCAR and RFLP markers tightly linked to nematode resistance genes introgressed from *Arachis cardenasii* into *Arachis hypogaea*. *Genome*, **39**(5): 836–845.

Grattapaglia D, Sederoff R. 1994. Genetic linkage maps of *Eucalyptus Grandis* and *Eucalytus urophy* using a Pseudo-testcross mapping strategy and RAPD marker. *Genetics*, **137**(4): 1121–1137.

Jermstad KD, Bassoni DL, Jech KS, Wheeler NC, Ritchie GA, Neale DB. 2003. Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas-fir III.QTL trait loci-by-environment interactions. *Genetics*, **165**(3): 1489–1506.

Li Ping, Fang Guiping, Sun Chengzhi. 1995. Wood characteristics of pulpwood. *Chemistry and Industry of Forest Products*, **15**(Suppl): 13–18.(in Chinese)

Meletto M. 1996. Development of a SCAR marker linked to the gene in common bean. *Genome*, **39**(6): 1216–1219.

Ohmori T. 1996. Molecular characterization of RAPD and SCAR markers linked to the Tm-1 locus in tomato. *Theoretical and Applied Genetics*, **92**(2): 151–156.

Shi Jiang, Xin Li, Tan Lin, Zheng Xueqin. 2007. Study on the RAPD specific band of Kava (*Piper, Piperaceae)* transferring to the SCAR molecular marker. *Acta Botanica Yunnanica*, **29**(4): 457–460. (in Chinese)

Tang Meiling, Kong Jin, Xu Xuefeng, Shen Yujie, Wang Yi, Li Tianzhong, Han Zhenghai. 2008. Screening of the AFLPM markers associated to the sex locus of *Vitis amurensis* Rupr and its conversion to SCAR marker. *Acta Horticulturae Sinica*, **35**(2): 195–200. (in Chinese)

Wu RL, Han HF, Hu JJ, Fang JJ, Li L, Li ML, Zeng ZB. 2000. An integrated genetic map of *Populus deltoids* based on amplified fragment length polymorphisms. *Theoretical and Applied Genetics*, **100**(8): 1249–1256

Xu Yunbi, Zhu Lihuang. 1994. The application of molecular quantitative genetics. Beijing: China Agriculture Press, pp.112–115. (in Chinese)

Xu Jichen, Zou Liangxing. 2002. Identification of molecular markers associated with rice root traits by correlation coefficient analysis. *Acta Genetica Sinica*, **29**(3): 245–249. (in Chinese)

Zeng Changying, Xu Fangsen, Meng Jinling, Wang Yunhua, Hu Chenxiao. 2006. How long the way from QTL to QTGs? *Hereditas*, **28**(9): 1191–1198.

Zhang Bo, Du Mingsheng, Huang Minren. 2003. Status quo and tendency in construction of forest frees' genetic linkage maps. *China Biotechnology*, **23**(4): 14–18.(in Chinese)

Zhan Yaguang, Zeng Fansuo. 2005. A method for DNA extraction from mature birch leaves rich in polysaccharid. *Journal of Northeast Forestry University*, **33**(3): 24–25.(in Chinese)

Zhang Zhiyi, Lin Shanzhi, Zhang Deqiang, Zhang Qian. 2002. Applications of modern molecular biotechnology to genetic improvement in forest trees. *Journal of Beijing Forestry University*, **23**(4): 250–261.(in Chinese)